



Title: COMPO

NS AND METHODS INVOLVING AN ESSEN L STAPHYLOCOCCUS
AUREUS GENE AND ITS ENCODED PROTEIN

Appl. No.: 10/025,222 Filing Date: 12/19/01 Applicant: Pelletier, et al.

Sheet 1 of 24

SEQ ID NO: 1

>STAAU_R009 nucleotide sequence: 1800

TTGCGAATAGATCAATCGATCATTAAATGAAATAAAAGATAAAACCGACATTTTACTGTTGTAAGTGA
ATATGTAAAATTAGAAAAGAGAGGACGCAATTATATAGGTTTGTGTCCTTTTCATGATGAAAAGACAC
CTTCATTTACAGTTTCTGAAGATAAAACAAATTTGTCAATTGTTTTGGTTGTAAAAAAGGTGGCAATGTT
TTTCAATTTACTCAAGAAATTAAAGACATATCATTGTTGAAGCGGTAAAGAATTAGGTGATAGAGT
TAATGTTGCTGTAGATATTGAGGCAACACAATCTAACTCAAATGTTCAAATTGCTTCTGATGATTTAC
AAATGATTGAAATGCATGAGTTAATACAAGAATTTTATTATTACGCTTTAAACAAAGACAGTCGAAGGC
GAACAAGCATTAAACGTACTTACAAGAACGTGGTTTTACAGATGCGCTTATTAAAGAGCGAGGCATTGG
CTTTGCACCCGATAGCTCACATTTTGTGTCATGATTTTCTTCAAAAAAAGGGTTACGATATTGAATTAG
CATATGAAGCCGGATTATTATCACGTAACGAAGAAAATTTAGTTATTACGATAGATTTCGAAATCGT
ATTATGTTTCTTTGAAAAATGCGCAAGGAAGAAATGTTGGATATTCAGGTGCAACATATACCGGTCA
AGAACCAAAATACTTAAATAGTCCTGAAACACCTATCTTTCAAAAAAGAAAGTTGTTATACAACCTTAG
ATAAAGCGCGTAAATCAATTAGAAAATTAGATGAAATCGTATTACTAGAAGGTTTTATGGATGTTATA
AAATCTGATACTGCTGGCTTGAAAAACGTTGTTGCAACAATGGGTACACAGTTGTCAGATGAACATAT
TACTTTTATACGAAAGTTAACATCAAATATAACATTAATGTTTGATGGGGATTTTGCGGGTAGTGAAG
CAACACTTAAACAGGTCAAAATTTGTTACAGCAAGGGCTAAATGTATTTGTTATACAATTGCCATCA
GGCATGGATCCGGATGAATACATTGGTAAGTATGGCAACGATGCATTTACTGCTTTTGTA AAAAATGA
CAAAAAGTCATTTGCACATTATAAAGTGAGTATATTA AAAAGATGAAATTGCACATAATGACCTTTCAT
ATGAACGTTATTTGAAAGAACTAAGTCATGATATTTGCTTATGAAATCATCGATTTTGCAACAAAAG
GCTTTAAATGATGTTGCACCATTTTTCAATGTTAGTCCTGAGCAATTAGCTAACGAAATACAATTCAA
TCAAGCACCGCAATTATTATCCAGAAGATGAGTATGGCGGTTACATTGAACCTGAGCCAATTGGTA
TGGCACAATTTGACAATTTGAGCCGTCAAGAAAAAGCGAGCGAGCATTTTTTAAACATTTAATGAGA
GATAAAGATACATTTTTTAAATTATTATGAAAGTGTTGATAAGGATAACTTCACAAATCAGCATTTTAA
ATATGTATTCGAAGTCTTACATGATTTTTATGCGGAAAATGATCAATATAATATCAGTGATGCTGTGC
AGTATGTTAATTCAAATGAGTTGAGAGAAACACTAATTAGCTTAGAACAATATAATTTGAATGACGAA
CCATATGAAAATGAAATTGATGATTATGTCAATGTTATTAATGAAAAAGGACAAGAAACAATTGAGTC
ATTGAATCATAAATTAAGGGAAGCTACAAGGATTGGCGATGTAGAATTACAAAAATACTATTTACAGC
AAATTGTTGCTAAGAATAAAGAACGCATGTAG

SEQ ID NO: 2

>STAAU_R009 amino acid sequence : 599

LRIDQSIINEIKDKTDILDVSEYVKLEKRGRNYIGLCPFHDEKTPSFTVSEDKQICHCFGCKKGGNV
FQFTQEIKDISFVEAVKELGDRVNVAVDIEATQSNSNVQIASDDLQMIEMHELIQEFYYYALTKTVEG
EQALTYLQERGFTDALIKERGIGFAPDSSHFCDFLQKKGYDIELAYEAGLLSRNEENFSYYDRFRNR
IMFPLKNAQGRIVGYSGRITYTGQEPKYLNSPETPIFQKRKLLYNLDKARKSIRKLDEIVLLEGFMDVI
KSDTAGLKNVATMGTOQLSDEHITFIRKLTSNITLMFDGDFAGSEATLKTGQNLLQOGLNVFVIQLPS
GMDPDEYIGKYGNDAFTAFVKNDKKSFAHYKVSILKDEIAHNDLSYERYLKELSHDISLMKSSILQOK
ALNDVAPFFNVSPQLANEIQFNQAPANYYPEDEYGGYIEPEPIGMAQFDNLSRQEKAERAFLKHLMR
DKDTFLNYYESVDKNFTNQHFKYVFEVLHDFYAENDQYNISDAVQYVNSNELRETLISLEQYNLNDE
PYENEIDDYVNVINEKGQETIESLNHKLREATRIGDVELQKYYLQQIVAKNKERM

FIGURE 2

SEQ ID NO: 3

>96ORF078 nucleotide sequence

ATGAATATAATGCAATTCAAAAGCTTATTGAAATCGATGTATGAAGAGACAAAGCAAAGC
GACCCGATTGTAGCAAATGTATATATCGAGACTGGTTGGGCGGTCAATAGATTGTTGGAC
AATAACGAGTTATCGCCTTTCGATGATTACGACAGAGTTGAAAAGAAAATCATGAATGAA
ATCAACTGGAAGAAAACACACATTAAGGAGTGTTAA

SEQ ID NO: 4

>96ORF078 amino acid sequence 96_NT|10148-10363|

MNIMQFKSLLKSMYEETKQSDPIVANVYIETGWAVNRLLDNNELSPFDDYDRVEKKIMNE
INWKKTHIEC

FIGURE 3

A.

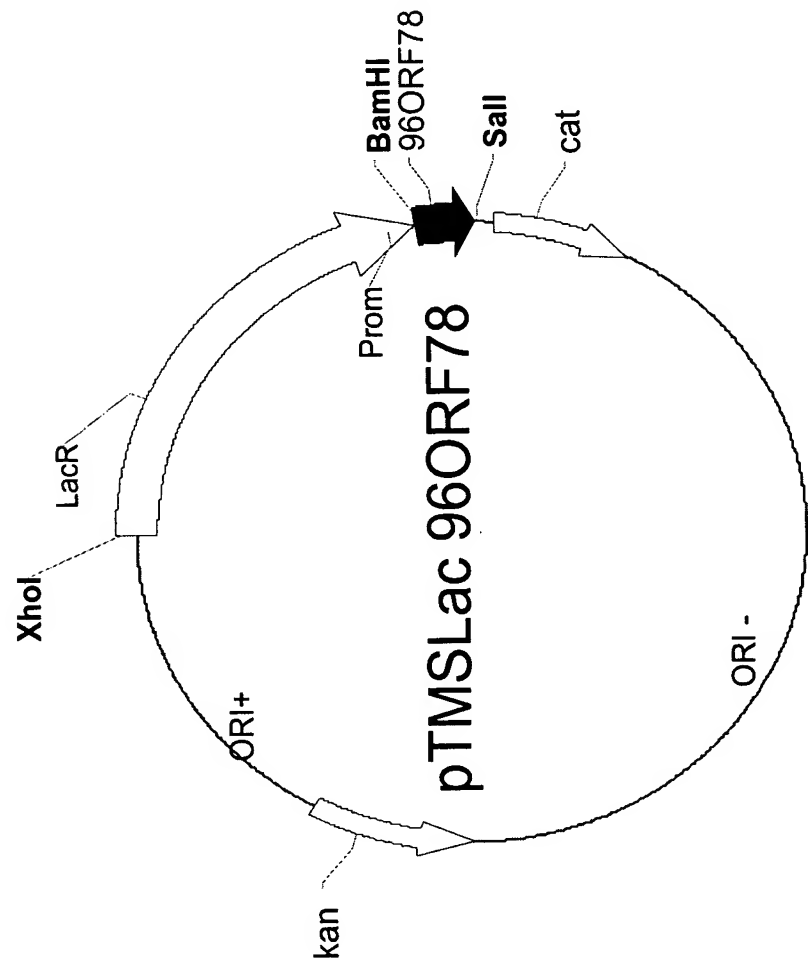


FIGURE 3 B.

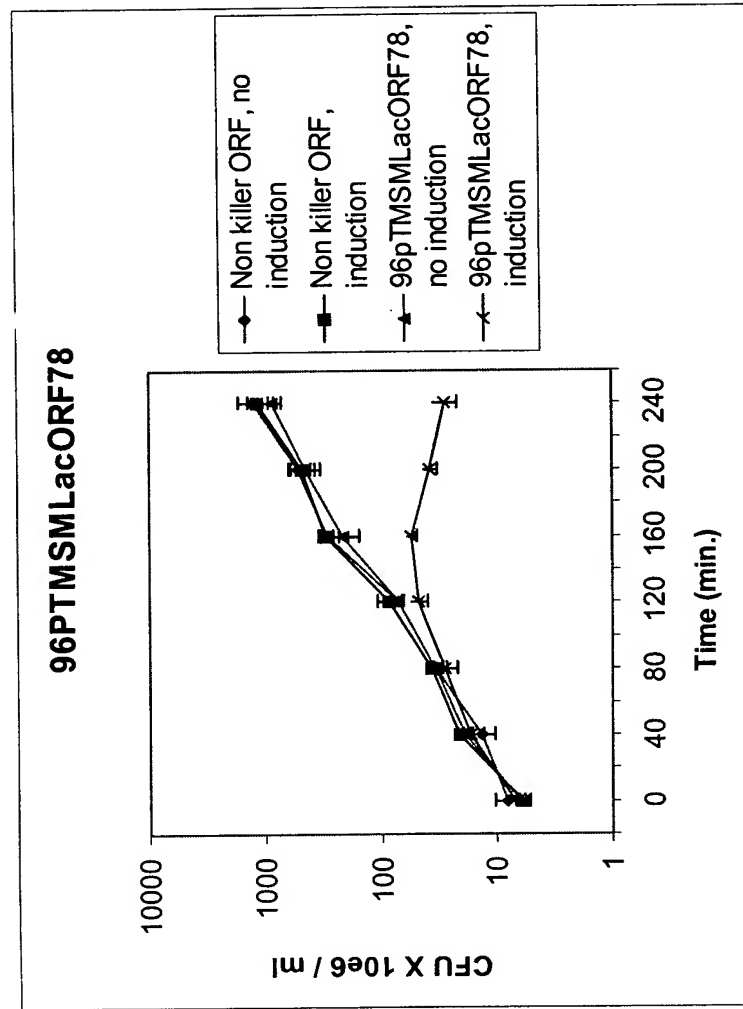
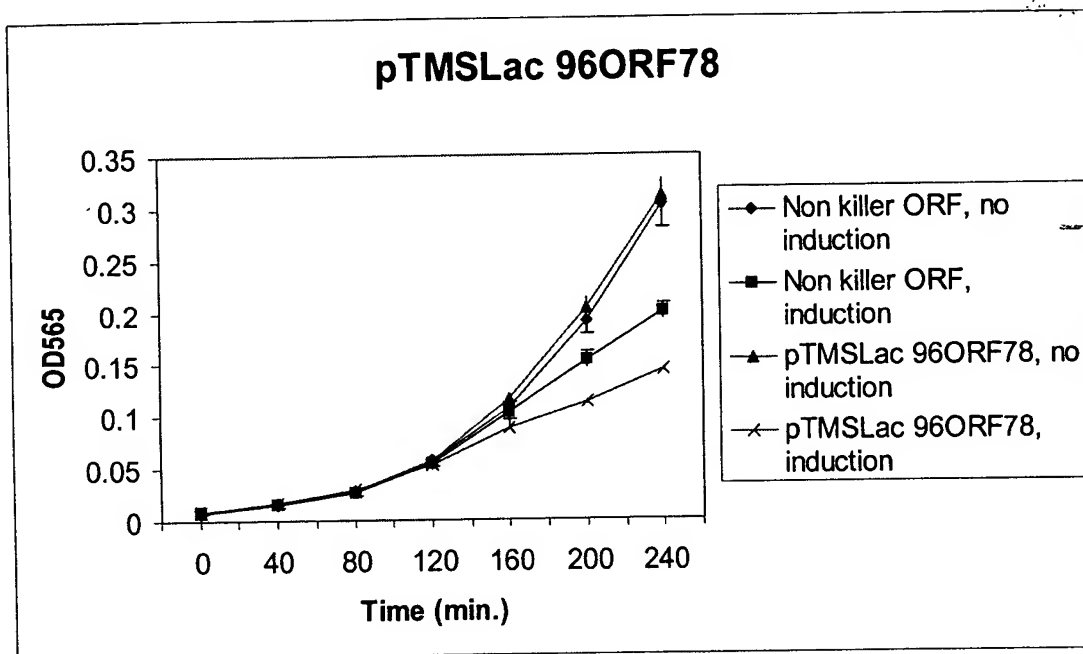


FIGURE 3 C.



D.

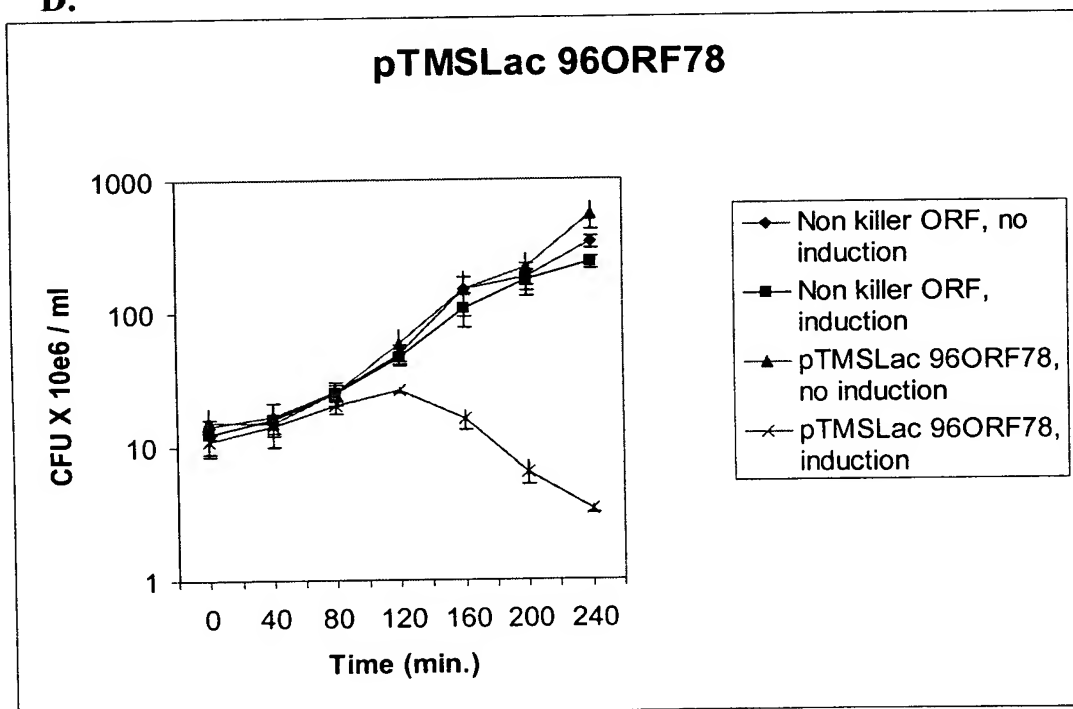


FIGURE 4

A. GST/96ORF78

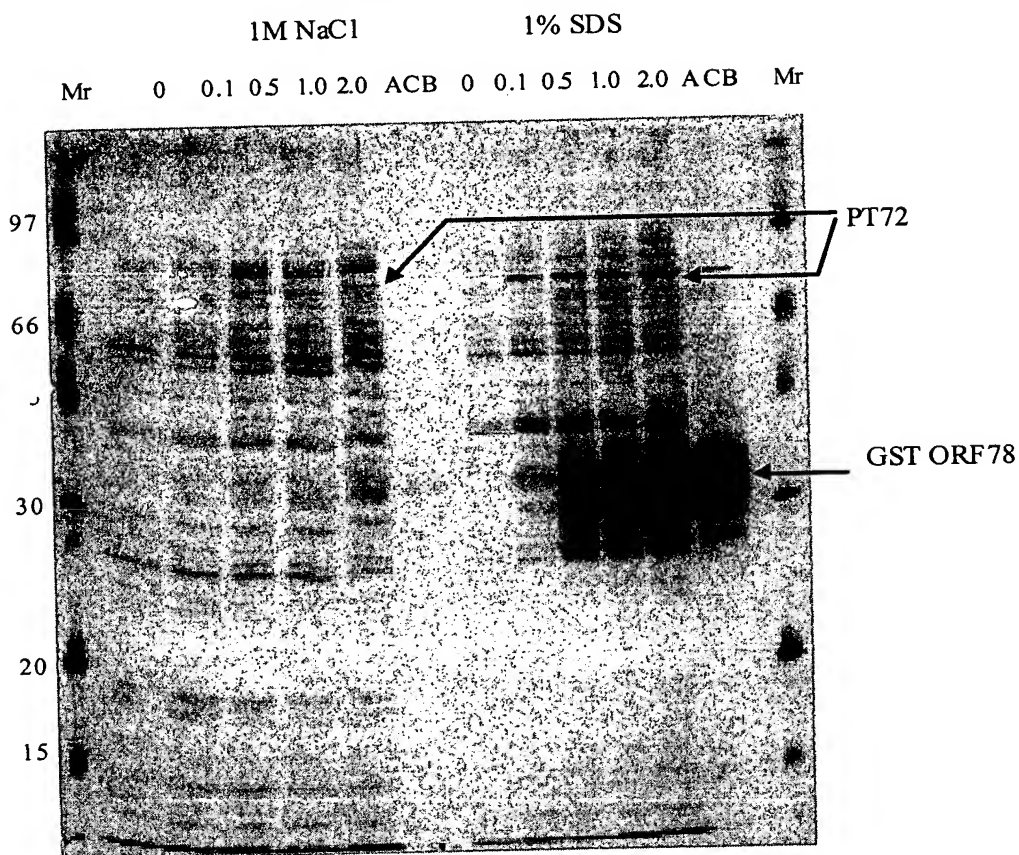


FIGURE 4

B. GST

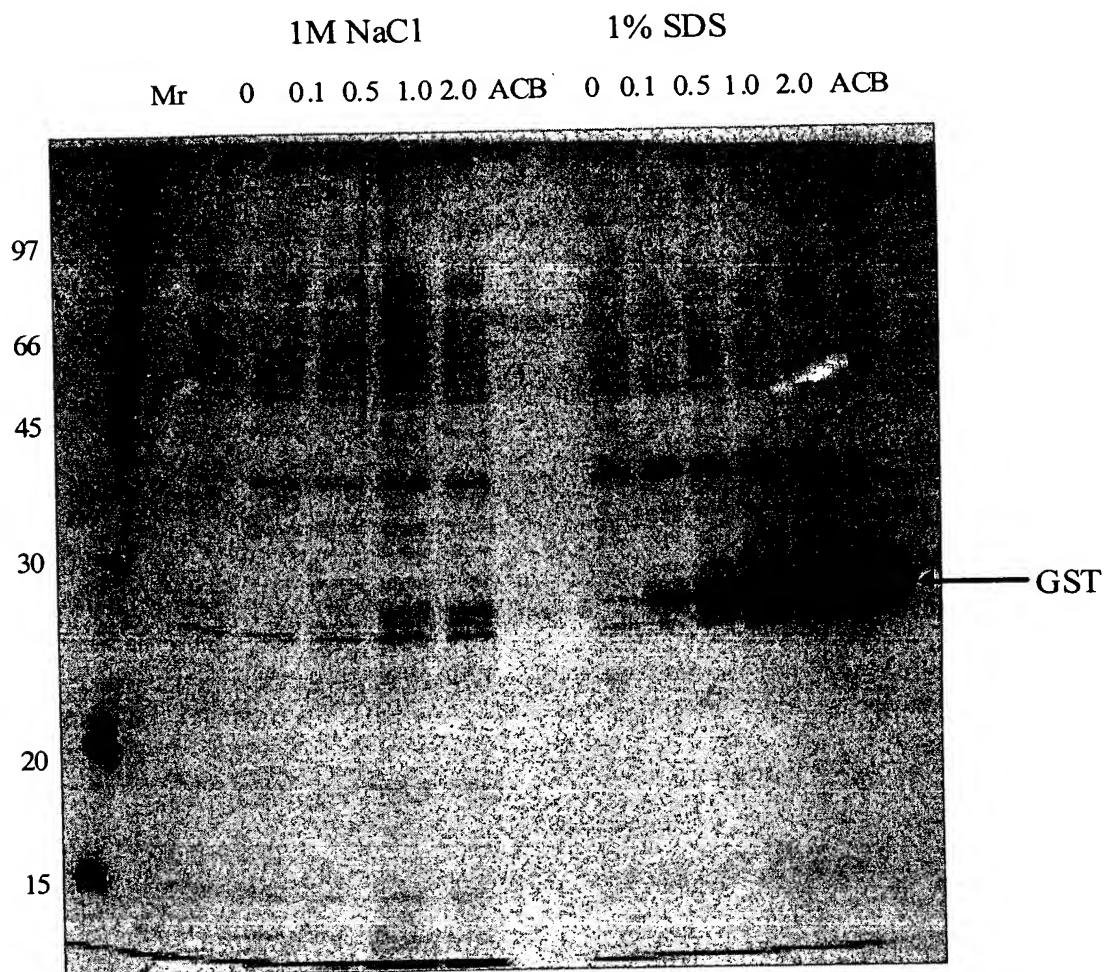


FIGURE 5 96ORF78 (GST removed)

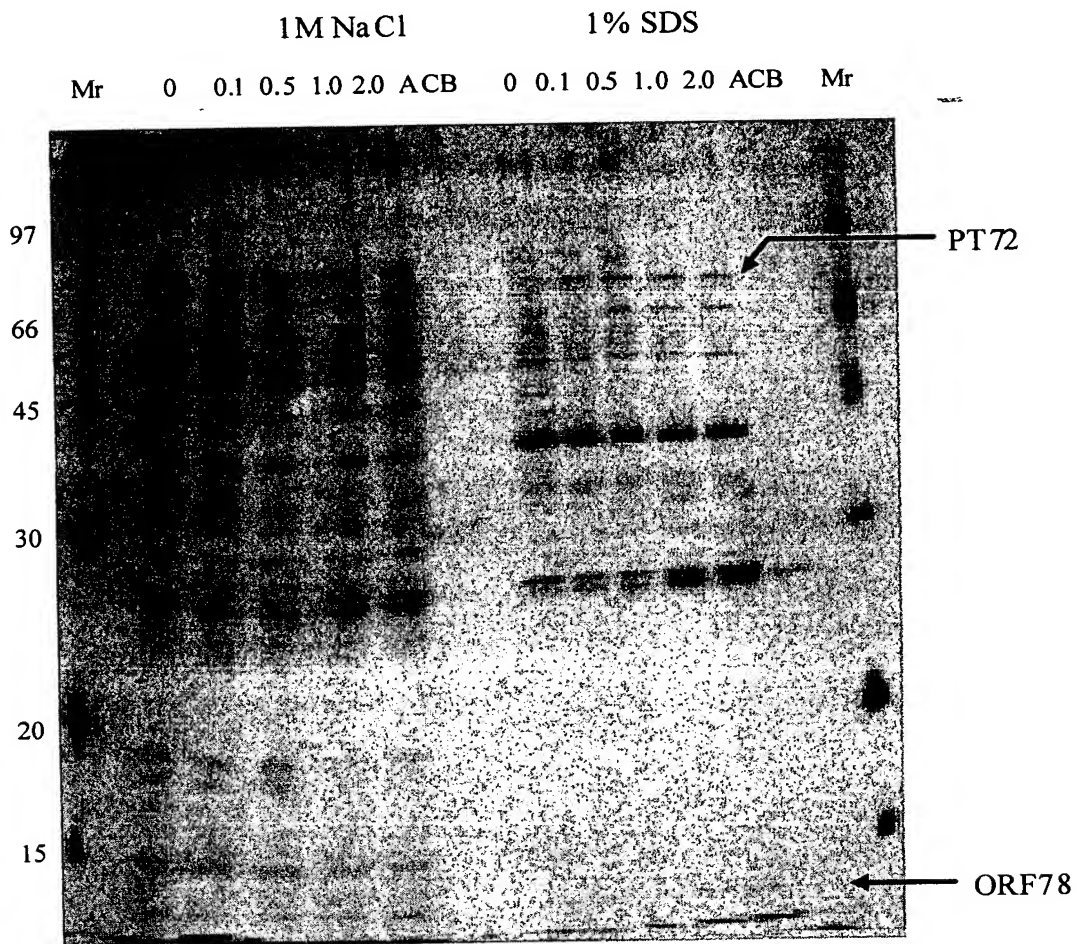
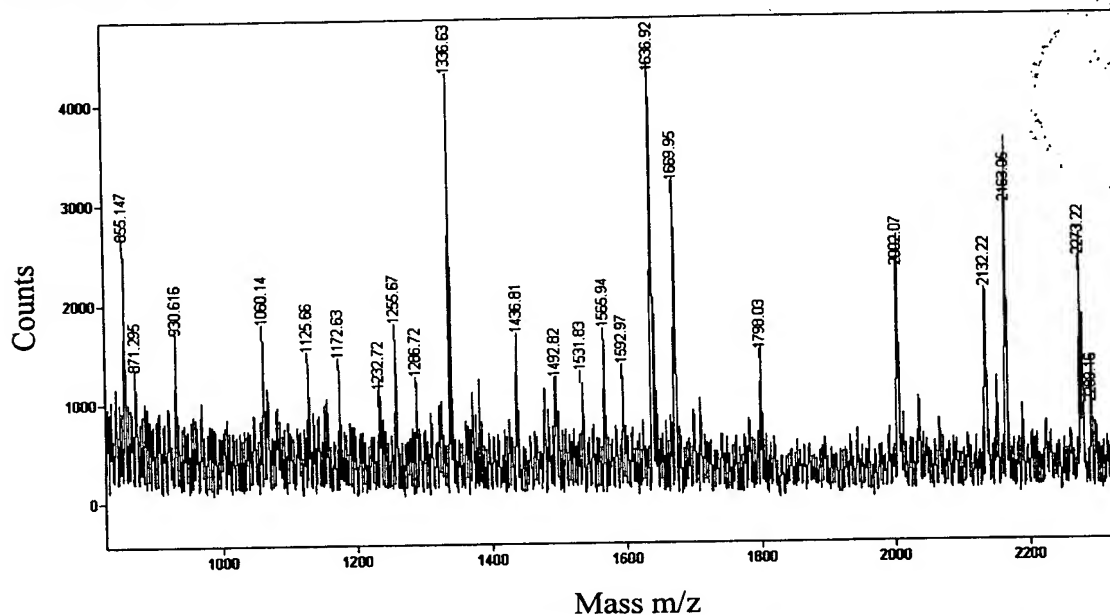


FIGURE 6



Details for rank 1 candidate in search 20000915095311-0166-127000000001

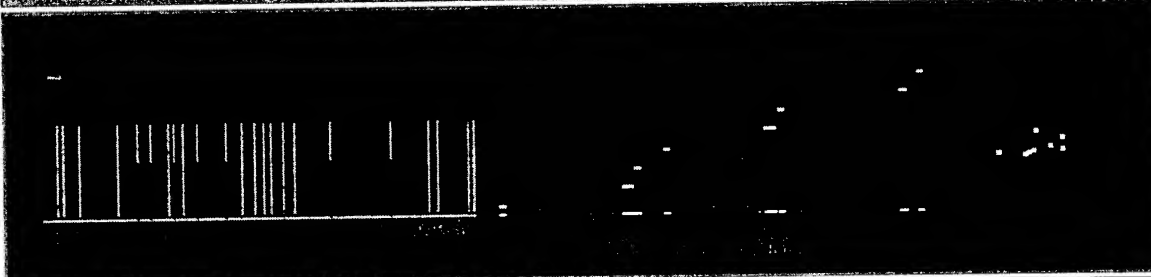
ouContig238: 82020 to 831822 Frame 1301 on

Sample ID: 10178.72 [Pass10]

Measured peptides: 23

Matched peptides: 7

Min. sequence coverage: 14%



Measured Avg/Computed Error Residues Missed

Mass(M) Mono Mass (Da) Start To Cut Peptide sequence

1124.649	M	1124.622	0.027	588	596	0	YYLQQIVAK
1171.623	M	1171.590	0.033	394	403	0	ELSHDISLMK
1171.623	M	1171.644	-0.021	5	14	0	IDQSIINEIK
1254.661	M	1254.620	0.041	563	573	0	GQETIESLNHK
1335.627	M	1335.536	0.091	193	202	0	NEENFSYYDR
1435.802	M	1435.734	0.068	233	244	0	YLNSPETPIFK
1797.019	M	1796.930	0.089	177	192	1	KGYDIELAYEAGLLSR
2001.060	M	2001.016	0.044	374	390	1	VSILKDEIAHNDLSYER

Unmatched Monoisotopic Masses:

855.147 871.295 930.616 1060.140 1232.717 1286.716 1492.822 1531.830 1565.943 1592.967
1636.921 1669.948 2132.217 2163.048 2273.222 2289.164

FIGURE 7A

A. Pfam HMM search results

Model Seq-from Seq-to HMM-from HMM-to Score E-value			
zf-CHC2	3	100	1 98 198.4 1.1e-55
Toprim	260	339	1 151 71.9 1.3e-17

Alignments of top-scoring domains:

zf-CHC2:
domain 1 of 1, from 3 to 100: score 198.4, E = 1.1e-55
*->ipeesIdeLknriDIVdvisYVvKkGrnYkgLCPFHdEKTSPFs
i++++I+e+k++++DI d++seYVvK+K+GrnY+gLCPFHdEKTSPF+
gi|133988 3 IDQSIINEIKDKTDILDVSEYVVKLEKGRNYYIGLCPFHdEKTSPFT 49

VspeKqfYhCFGCGagGdaIkFlmkyeklsFvEavekLAdragidlpyek
Vs +Kq+ hCFGc++g+++ F ++++++sFvEav++L dr+++ +++e
gi|133988 50 VSEDKQICHCFGCKKGGNVFQFTQEIKDISFVEAVKELGDRVNVAVDIEA 99

g<-*
+
gi|133988 100 T 100

Title: COMPOSITIONS AND METHODS INVOLVING AN ESSENTIAL STAPHYLOCOCCUS
 AUREUS GENE AND ITS ENCODED PROTEIN
 Appl. No.: 10/025,222 Filing Date: 12/19/01 Applicant: Pelletier, et al.
 Sheet 12 of 24

Toprim: domain 1 of 1, from 260 to 339: score 71.9, E = 1.3e-17

gi 133988	260	DEIVLLEGFMVDVIKSDTAGL-----KNVVA	289
		+++++ Eg++d+++ a+ +nvva++G+	
		*->kvliiVEgpsdakalalgkpskrkiyyelpggkdgngvvaslGhlv	
gi 133988	290	-----LSDEHI	295
		dlptpegyddkykwlwlpivdvkkgefepyqiefdqkcskklkkeql	
		l+++e++	
gi 133988	296	TFIRKLTSN---ITLMFDGDFAGSEATLKGTQNLLQQCL	335
		+++kl+++ + l++D+D +G +++++k + l+ +q +v++	
		kllklklakkadevilatDpDreGeaiawkllellkpygpveledkdvrr	
gi 133988	336	IQLP	339
		iflp<-*	
		i+lp	

FIGURE 7B

B. Optimal global alignment of amino acid sequences

Sequence 1 sp|O05338|PRIM_STAAU DNA PRIMASE (EC 2.7.7.-) - S. aureus. (572 letters)
Sequence 2 STAAU_R009 STAAU_R009_NT|1-1800| (599 letters)

Identical: 560/605 (92%), Similar: 564/605 (93%), Gap: 39/605 (6%)

seq1	1	M-----	IGLCPFHDEKTPSFTVSEDKQICHCF	27
		:		
seq2	1	LRIDQSIINEIKDKTDILDVSEYVKLEKGRNYIGLCPFHDEKTPSFTVSEDKQICHCF		60
		:		
seq1	28	GCKKGGNVFQFTQEIKDISFVEAVKELGDRNVAVDIEATQSNSNVQIASDDLQMIEMHE		87
seq2	61	GCKKGGNVFQFTQEIKDISFVEAVKELGDRNVAVDIEATQSNSNVQIASDDLQMIEMHE		120
seq1	88	LIQEFYYYALTKTVEGEQALTYLQERGFTDALIKERGIGFAPDSSHCHDFLQKKGYDIE		147
seq2	121	LIQEFYYYALTKTVEGEQALTYLQERGFTDALIKERGIGFAPDSSHCHDFLQKKGYDIE		180
seq1	148	LAYEAGLLSRNEENFSYYDRFRNRIMFPLKNAQGRIVGYSGRITYTGQEPKYLNSPETPIF		207
seq2	181	LAYEAGLLSRNEENFSYYDRFRNRIMFPLKNAQGRIVGYSGRITYTGQEPKYLNSPETPIF		240
seq1	208	QKRKLLYNLDKARKSIRKLDEIVLLEGFMDVIKSDTAGLKNVATMGTLSDHEITFIRK		267
seq2	241	QKRKLLYNLDKARKSIRKLDEIVLLEGFMDVIKSDTAGLKNVATMGTLSDHEITFIRK		300
seq1	268	LTSNITLMFDGDFAGSEATLKTGQHLLQQGLNVFVIQLPSGMDPDEYIGKYGNDAFTTFV		327
seq2	301	LTSNITLMFDGDFAGSEATLKTGQNLQQGLNVFVIQLPSGMDPDEYIGKYGNDAFTAFV		360

seq1	328	KNDKKSF	AHYKVS	ILKDEIA	HNDSL	YERYL	KELSHD	ISLMK	SSILQ	QKAI	NDVAP	FFNV	387
seq2	361	KNDKKSF	AHYKVS	ILKDEIA	HNDSL	YERYL	KELSHD	ISLMK	SSILQ	QKAI	NDVAP	FFNV	420
seq1	388	PEQLANE	IQFNQ	APANY	PEDEY	GGYDE	YGGYI	EP	IGMAQ	FDNL	SRREK	AERAF	447
seq2	421	PEQLANE	IQFNQ	APANY	PE	-----	DEYGG	YIEP	IGMAQ	FDNL	SRQEK	AERAF	474
seq1	448	MRDKDT	FLNYY	ESVDK	NFTNQ	HFKYV	FEVLH	DFYA	ENDQ	YNISD	AVQY	VNSNEL	507
seq2	475	MRDKDT	FLNYY	ESVDK	NFTNQ	HFKYV	FEVLH	DFYA	ENDQ	YNISD	AVQY	VNSNEL	534
seq1	508	SLEQYN	LNGEP	YENEI	DDYVN	VINEK	GQETI	ESLN	HKLRE	ATRIGD	VELQ	KYYLQ	567
seq2	535	SLEQYN	LNDPE	YENEI	DDYVN	VINEK	GQETI	ESLN	HKLRE	ATRIGD	VELQ	KYYLQ	594
seq1	568	NKERM											
seq2	595	NKERM											


```
seq2 358 AGEAGCRRPLVAFKMYLR--RGKNLQHEGERLLRYIDEALREIGKLSSPVEQDYVLRQL 414
seq1 414 APFFNVSPQLANEIQFNQAPANYYPEDGYGIEPEPIGMAQFDNLSRQEKAEAFLLKH 473
      | | : | | : : | | | | | | | | : | | | | |
seq2 415 AEEFSLSLSALHEQLSRQERTKPREAPDGETARP----MLAKKLLPAFQNAERLLLAH 470
474 LMRDKDTFLNYYESVDKNFTNQHFYVFEVLHDFYAENDQYNI SDAVQYVNSNELRETL 533
      : | | : | | : : : : : : : : : : : : : : : :
seq2 471 MMRSRDVALVQERIG-GRFNIEEHRALAAAYIAFYEEGHEADPGALLSRI-PGELQPLA 528
534 ISLEQYNLNDEPYENEIDDDYV-NVINEKGQETIESLNHKLREATRIGDVELQKYVLRQIV 592
      : : | : | | : : | : : : : : : : : : : : : : : :
seq2 529 SDVSLILLIADDVSEQEELEYIRHVLRPKWMLKVKQEKTAEARRKDFLTAARIAKEMI 588
593 AKNK--ERM 599
      |
seq2 589 EMKKMLSSS 597
```



```
seq1 359 FVKNDKKSFAHYKVSIL-KDEIAHNDLSYERYLKELSHDISLMKSSILQQKALNDVAPFF 417
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
seq2 361 DIIDASVTVMFAFKMQYFRKGKNSDEGDRLAYIKDVLKEISTLSGSLEQEVYVKQLASEF 420

seq1 418 NVSPEQLANEIQFNQAPANYYPEDEYGGYIEPEPIGMAQFDNLSRQ-----EKAERAF 470
      ::||| | | : | | | | : || | | | | | | | | | | | |
seq2 421 SLSQESLTE--QLSVFSKQNKPADNSG----ETKTRRAHLTTKARQKRLRPAYENAEERLL 474

seq1 471 LKHLMRDKDTFLNYYESVDKDNFTNQHFKYVFEVLHDFYAENDQYNISDAVQYVNSNELR 530
      | | ::|| : | : : : : : : : : : : : : : : : : : :
seq2 475 LAHMLRDRSVIKKVIDRVGFQFNIDEH-RALAAYLAFYEEGAELTPQHLMARVTTDDHIS 533

seq1 531 ETLISLEQYNLNDEPYENEIDYV-NVINEKGQETIESLNHKLREATRIGDVELQKYYLQ 589
      : | : : | | | | : ||| | : : : : : : : : | |
seq2 534 QLLSDILMLQVNQELSEAEISDYVKKVLNQPNWSMIKEKEAERAERQKDFLRAASLAQ 593

seq1 590 QIVAKNKERM 599
      : || | :
seq2 594 EIVTLNRSRK 603
```



ALPHEIIS GENE AND ITS ENCODED PROTEIN

Sheet 19 of 24

(599 letters)

Category	Count	Percentage
Category 1	58	9%
Category 2	619	97%

1: 170/619 (2/8), similar: 254/615 (1/8) ref.

[illegible][illegible]

61 CFCGAGHGNATDFEINTIDANEEI VFAVZEEZEE

19 HELIQEYYALTKTVEGEQALTYLQERGFIDALINERGICITATDSSMCHBTF

20 MDGLNTFYQQSLQQPV-ATSARQYLEKRGLSHEVIARFAIGFAFFG---

79 TE---LAYEAGLLSRNEENFSYDRFRNRIMFPLKNAQGRIVGYSGRITYTGQEPKYLNSI

70 PENPQSLIDAGMTVTNDQGPS-YDRFRERVMPPIRDKRGRVIGGGRVLGNDTPKYLNSP

----- * ***** DEIVY I EGEMDVTKSDTAGI.KNVVATMGTOLSDEHI

[illegible][illegible]

667 JPTNNTANOTOC

353 NDAF TAF VKNDKSF AF IN SVI FØDETA FNDS IETAT FØRST

355 KEAFEARMEQAMPLSAFLFNSLMPQV-----DLS-----IPDGRKATSIATFFHSQVIST

Title: COMPOSITIONS AND METHODS INVOLVING AN ESSENTIAL STAPHYLOCOCCUS
AUREUS GENE AND ITS ENCODED PROTEIN
Appl. No.: 10/025,222 Filing Date: 12/19/01 Applicant: Pelletier, et al.
Sheet 20 of 24

seq1	413	VAPFFNVSP	EQLANEIQF--NQAPANYYP	EDEYGGYIEPEPIGMAQFDNLSRQ	KAERAF	470
		:	:	:	:	:
		:	:	:	:	:
seq2	407	LRIYLR---	QELGNKLGILDDSQLERLMPKAAESG	VRPVP-----QLKRTTMRIL		454
		:	:	:	:	:
		:	:	:	:	:
seq1	471	LKHLMRD	KD--TFLNYYESVDKNFTNQHF	KYVFEVLHDFYAENDQYNI	SDAVQYVNSNE	528
		:	:	:	:	:
		:	:	:	:	:
seq2	455	IGLLVQN	PELATLVPPLENLDENKLPGLG--	LFRELVNTCLSQPGLTTGQLLEHYRGTN		511
		:	:	:	:	:
		:	:	:	:	:
seq1	529	LRETLIS	LEQY-----NLNDEPYENEID	DDYVNVINEKGQETIESLNHKLREATR--	IGD	580
		:	:	:	:	:
		:	:	:	:	:
seq2	512	NAATLEK	LSMWDDIADKNIAEQTFTDSLNHMF	DSLLELRQEELIA-----RERTHGLSNE		566
		:	:	:	:	:
		:	:	:	:	:
seq1	581	VELQKYY	LQQIVAKNKERM			599
		:	:	:	:	:
		:	:	:	:	:
seq2	567	ERLELWT	LNQELAKK----			581

FIGURE 8

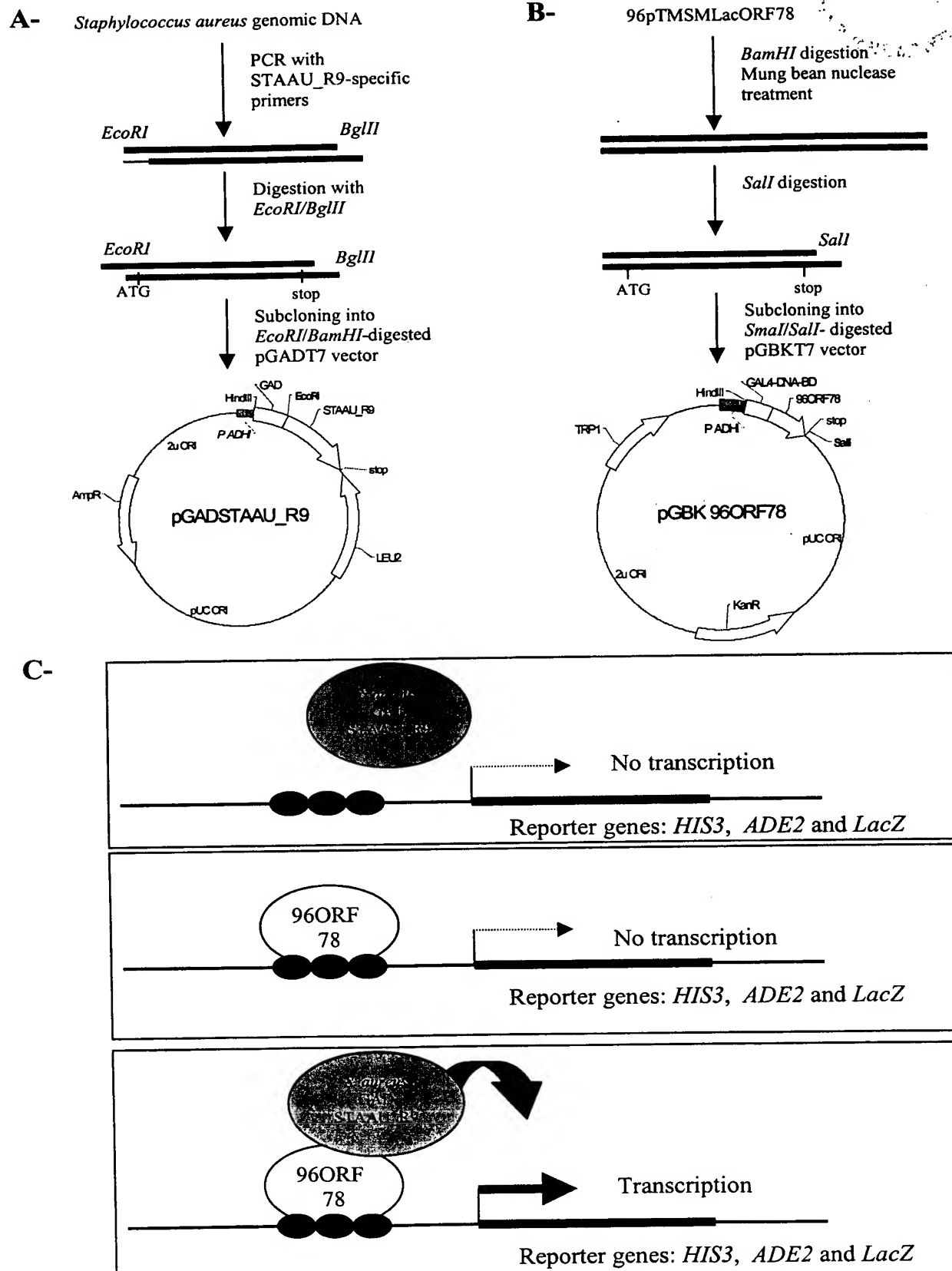
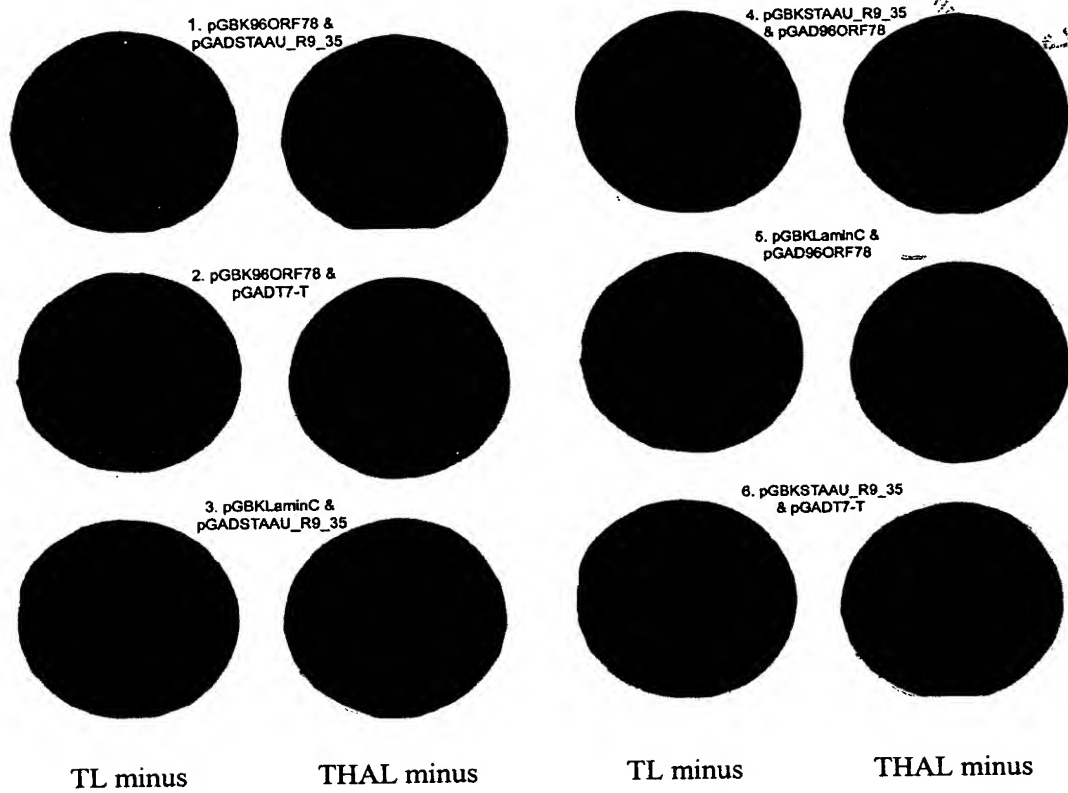
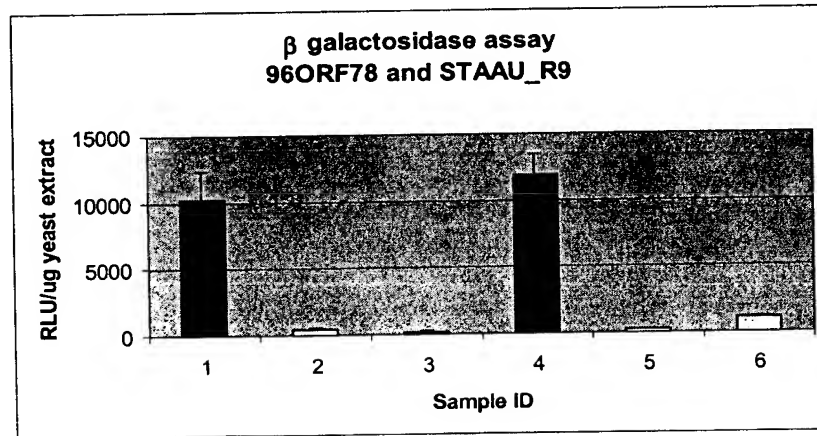


FIGURE 9

A.



B.



Sample	Yeast co-transformants		RLU (avg)	SD
	Plasmid 1	Plasmid 2		
1	pGBK 96ORF78	pGADSTAAU_R9_35	10, 250	2, 080
2	pGBK 96ORF78	pGADT7-T	475	38
3	pGBK Lamin C	pGADSTAAU_R9_35	224	10
4	pGBKSTAAU_R9_35	pGAD 96ORF78	11, 935	1, 477
5	pGBK Lamin C	pGAD 96ORF78	243	12
6	pGBKSTAAU_R9_35	pGADT7-T	1, 121	37

FIGURE 10

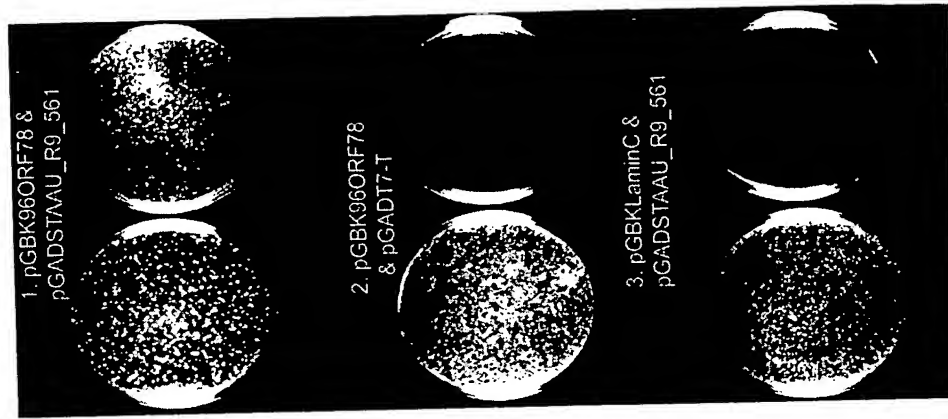
A. Fragments of STAAU_R9

B.

Interaction with
96ORF78

1	599	Yes
35	599	Yes
35	342	No
229	402	No
229	599	Yes
380	599	Yes
380	449	No
380	490	No
380	530	No
380	561	No
449	599	Yes
490	599	Yes
530	599	Yes
561	599	Yes

C.



SEQ ID NO: 6 STAAU_R9_561-599
AMINO ACID SEQUENCE

GQETIESLNHKLREATRIGDVELQKYLLQIVAKNKERM

SEQ ID NO: 5 STAAU_R9_1683-1800
NUCLEIC ACID SEQUENCE

GGACAAGAAACAATTGAGTCATTGAATCATAAATTAAGGGAAGCTACAAGGATTGGCGATGTA
GAATTACAAAAATACTATTACAGCAAAATTGTGCTAAGAAATAAAGAACGCATGTAG

FIGURE 11

A.

	Primer name	Primer sequence	Restriction site
SEQ ID NO: 8	R9_5E3	5'-ccggaattcTTGCGAATAGATCAATCG-3'	EcoRI
SEQ ID NO: 9	R9_3BG	5'-ggaagatctCTACATGCGTTCTTTATTC-3'	BglII
SEQ ID NO: 10	R9_5E	5'-ccggaattcATGATAGTTTGTGTCT-3'	EcoRI
SEQ ID NO: 11	R9_5E1	5'-ccggaattcCCAAAATACCTAAATAGTCC-3'	EcoRI
SEQ ID NO: 12	R9_5E2	5'-ccggaattcGCACATAATGACCTTTCA-3'	EcoRI
SEQ ID NO: 13	R9_342R	5'-cgcgatccATGCCTGATGGCAATTG-3'	BamHI
SEQ ID NO: 14	R9_402R	5'-ccatcgatGATTTTCATAAGCGAAATATC-3'	ClaI
SEQ ID NO: 15	R9_449F	5'-ccggaattcCCTGAGCCAATTGGTATGGC-3'	EcoRI
SEQ ID NO: 16	R9_449R	5'-cgcgatccctaAGGTTCAATGTAACCGCC-3'	BamHI
SEQ ID NO: 17	R9_490F	5'-ccggaattcAAGGATAACTTCACAAATCAG-3'	EcoRI
SEQ ID NO: 18	R9_490R	5'-cgcgatccctaCTTATCAACACTTTCATAATA-3'	BamHI
SEQ ID NO: 19	R9_530F	5'-ccggaattcAGAGAAACACTAATTAGCTTA-3'	EcoRI
SEQ ID NO: 20	R9_530R	5'-cgcgatccctaTCTCAACTCATTTGAATTAAC-3'	BamHI
SEQ ID NO: 21	R9_561F	5'-ccggaattcGGACAAGAAACAATTGAGTC-3'	EcoRI
SEQ ID NO: 7	R9_561R	5'-cgcgatccctaTCCTTTTTCATTAATAACATTG-3'	BamHI

B.

Cloning of SEQ ID NO: 2 amino acid fragments	Sense primer	Antisense primer
1-599	R9_5E3	R9_3BG
35-599	R9_5E	R9_3BG
35-342	R9_5E	R9_342R
229-402	R9_5E1	R9_402R
229-599	R9_5E1	R9_3BG
380-599	R9_5E2	R9_3BG
380-449	R9_5E2	R9_449R
380-490	R9_5E2	R9_490R
380-530	R9_5E2	R9_530R
380-561	R9_5E2	R9_561R
449-599	R9_449F	R9_3BG
490-599	R9_490F	R9_3BG
530-599	R9_530F	R9_3BG
561-599	R9_561F	R9_3BG